The OECD Program to Validate the Rat Uterotrophic Bioassay: An Overview

William Owens[1] and Herman B.W.M. Koëter[2]

[1]The Procter & Gamble Company, Cincinnati, Ohio, USA; [2]Environment, Health and Safety Division, Organisation for Economic Co-operation and Development, Paris, France

The Organisation for Economic Co-operation and Development has undertaken an international validation program for the rodent uterotrophic bioassay. This validation program comprised two major parts. The first part was the development of a detailed background review document compiling the existing data on the bioassay's history, the molecular and physiologic basis for the bioassay's mechanistic relevance to detect estrogen agonists and antagonists, a review of important bioassay protocol parameters, and a review of the data generated by *in vitro* assays, previous uterotrophic bioassays, and developmental and reproductive assays to assess and support the overall predictivity of the uterotrophic bioassay. The second part was an extensive multiyear effort managed by a validation management group to demonstrate the operating characteristics of four protocols. The effort was conducted in two phases. The phase 1 results with the reference agonist ethinyl estradiol (EE) and antagonist ZM 189,154 has been published previously. This *Environmental Health Perspectives* mini-monograph is devoted to the phase 2 work using five weak estrogen agonists, bisphenol A, genistein, methoxychlor, nonylphenol, and *o,p′*-DDT, as well as the negative substance dibutylphthalate. These data show that all protocols successfully detected increases in uterine weights when a sufficient dose level of the weak agonists was administered, whether the substances were known or provided as coded doses to the laboratory. The data with both the reference EE and all five weak agonists are reproducible over time and under a variety of different experimental conditions (e.g., animal strain, diet, housing, bedding, vehicle, animal age). In conclusion, all protocols now have sufficient data to support their validity. *Key words:* endocrine disruption, estrogen, rat uterus, uterotrophic, validation. *Environ Health Perspect* 111:1527–1529 (2003). doi:10.1289/ehp.6413 available via *http://dx.doi.org/* [Online 8 August 2003]

The Organisation for Economic Co-operation and Development (OECD) began the development of new and revised guidelines for the screening and testing of potential endocrine disrupters in 1998 (OECD 1998a). One activity is the validation of the rodent uterotrophic bioassay, an *in vivo* screen to identify suspected estrogen agonists or antagonists of estrogen. A validation management group (VMG, mammalian) has managed the uterotrophic validation program and has reported to the Task Force on Endocrine Disrupters Testing and Assessment.

Organization of the Validation Program

OECD Test Guidelines are used in all its member countries for purposes of the regulatory assessment of chemicals, and the validation program was international in scope. This international nature introduced numerous challenges, including different languages, the logistics of chemical distribution to nine countries, and the collection and statistical analysis of the validation data. The validation program was conducted by a VMG that recognized and addressed these challenges. The VMG designed the program to be consistent with recognized validation principles (ECVAM 1995; ICCVAM 1997; OECD 1998b). The process from initial discussions to develop the protocol to the completion of the reports and the following manuscripts has taken 4 years.

The first challenge was that the protocols required flexibility, as the animal strains, animal husbandry supplies, vehicles, and reagents had to be widely and readily available among the OECD member countries. Any rigorous and detailed standardization of all of these variables would have constrained the future ability to widely and easily practice the uterotrophic bioassay in many of the OECD member countries. To acquire, store, and distribute the chemicals, a central repository was established to deal with international shipments, including different customs regulations and laboratory safety regulations. The repository was kindly funded by the chemical industry of Europe and Japan. The VMG authored and then distributed the draft protocols for comments to all national authorities and participating laboratories, including inquiries for any language ambiguities. To avoid language issues, laboratory procedures were videotaped, and videotapes were distributed to the technical staff of all the participating laboratories. A common electronic spreadsheet was constructed and distributed for comment so the data could be recorded and electronically transmitted to the independent statistician.

The VMG organized the validation program in two major phases. Phase 1 was conducted as a protocol demonstration with potent references including an estrogen agonist, 17α-ethinyl estradiol (EE), and antagonist, ZM 189,154 (Kanno et al. 2001; OECD 2001). Phase 2 was conducted with compounds with weak agonists that likely represented target compounds of regulatory interest. To assess the predictivity of the uterotrophic bioassay, the weak agonists were selected because of the availability of reproductive and developmental studies with a battery of estrogen-sensitive end points (Chapin et al. 1999; Tyl et al. 2002). The chemicals were also selected with a range of potencies and different metabolic and pharmacokinetic properties in mind. Kanno et al. (2003a, 2003b) provide additional details on the chemical selections.

The VMG also oversaw the statistical analyses, the development of OECD program reports, the production of the detailed background review document, and the publication of the manuscripts. We thank VMG participants for their dedicated efforts over the past 4 years and their contribution to the success of this program.

Detailed Background Review Document

An essential part of a validation program is the detailed background review document. The OECD document has been completed (OECD 2003) and published in the

This article is part of the mini-monograph "The OECD Validation of the Uterotrophic Bioassay."

Address correspondence to J.W. Owens, Central Product Safety, The Procter & Gamble Company, 11810 East Miami River Rd., Cincinnati, OH 45252 USA. Telephone: (513) 627-1385. Fax: (513) 627-1208. E-mail: owens.jw@pg.com

Address any correspondence or questions about the OECD Programme and OECD documents to H.B.W.M. Koëter, OECD Environment, Health and Safety Division, OECD, 2 rue André Pascal, 75775 Paris Cedex 16, France. Telephone: 33 1 45 24 9849. Fax: 33 1 45 24 1675. E-mail: herman.koeter@oecd.org

We acknowledge the dedicated efforts and work of the participating laboratories in generating these data, the support of the chemical repository for this validation by the European Chemical Industry Council and the Japanese Chemical Industry Association, the work of the OECD Endocrine Disrupters Testing and Assessment Validation Management Committee, and many others for their advice, counsel, time, and assistance.

The authors declare they have no conflict of interest.

Received 25 April 2003; accepted 30 April 2003.

peer-reviewed literature (Owens and Ashby 2002). The background document comprises several parts. The first part is the history of the development and use of the uterotrophic bioassay beginning in the 1930s to identify potent estrogens and to allow their purification and structural elucidation. In subsequent decades, the bioassay was modified as part of the drug discovery process for estrogen agonists and antagonists and has been combined with preceding receptor binding and reporter gene transcriptional assays.

The second part provides the essential background to support the relevance of the uterotrophic bioassay. Uterine growth is the result of a cascade of events beginning with the docking of an agonist ligand with the estrogen receptor-α and culminating in cell division and growth of uterine tissues. The background document details the extensive work in this area including estrogen receptor knockout mice and antagonists. This body of work supports that the uterotrophic bioassay and the use of uterine growth is directly mechanistically based and is not empirical or correlative in nature.

The third part provides a detailed and critical review of the protocol variations and parameters. This includes work with rats and mice, immature females and ovariectomized females (as both have the requisite low endogenous circulating estrogen levels), the dissection and processing of the uterus (where imbibed fluid is often removed by blotting before weighting), and the possibility of supporting measurements such as histology to measure changes in the cell height of the uterine epithelium.

The fourth part uses several compounds as examples for a review of estrogen receptor binding, *in vitro* reporter gene transcriptional assays, pharmacokinetic experiments, and a body of reproductive and developmental data. This portion of the review presents evidence that the uterotrophic bioassay results appear to be generally concordant with estrogen effects (e.g., acceleration of vaginal opening) in several reproductive and developmental assays. Where not concordant, other toxicities such as body weight losses occur at lower doses than indicated by the uterotrophic bioassay.

In conclusion, the detailed background review indicates that the uterotrophic bioassay is mechanistically relevant and that existing data indicate that the uterotrophic bioassay is predictive in practice.

Phase 2 Validation Studies

The phase 2 validation studies were designed to generate the data to assess the reliability and reproducibility characteristics of the uterotrophic bioassay with weak estrogen receptor agonists and a negative substance. Responses consistent with an estrogen mode of action by the test chemicals were measured as a statistically significant increase in uterine weight versus vehicle-treated controls. The weights of uteri containing luminal fluid (wet weight) and with the luminal fluid removed (blotted weight) were examined to determine the procedure that provided the most reliable data or whether either approach could be used.

Phase 1 of the validation of the rodent uterotrophic bioassay involved 19 laboratories and measured responses of the animals to a potent reference estrogen, EE. Despite differences in rat strains, vehicles, and diets used as well as different levels of experience among the participating laboratories, there was acceptable agreement among laboratories with respect to the doses at which statistically significant responses were obtained for a given protocol (Kanno et al. 2001; OECD 2001).

Phase 2 involved 20 laboratories and seven test substances. These included the reference estrogen EE and five weak estrogen agonists that were three or more orders of magnitude lower in estrogen receptor binding affinity than in EE: bisphenol A (BPA), 1,1,1-trichloro-2,2-bis(*o,p′*-chlorophenyl)methane or *o,p′*-DDT (DDT), genistein (GN), methoxychlor (MX), and nonylphenol (NP). The weak agonists were tested in dose–response experiments with up to five doses per chemical and were provided blind to the laboratories in coded single-dose experiments. The coded single-dose experiments also included one chemical (dibutylphthalate; DBP) as a negative chemical having no estrogenic binding affinity or biological activity in well-designed and well-conducted studies.

All protocols were able to detect each of the five weak estrogen agonists, provided sufficient dose levels were administered. In phase 1, the minimal effective dose (MED; the lowest dose concentration in a series achieving statistical significance) for EE in protocol A by oral gavage was 0.3–1.0 mg/kg/day and in protocols B, C, and D by sc injection was 0.1–0.3 mg/kg/day. In phase 2, the MEDs of the weak agonists by oral gavage ranged from less than 20 mg/kg/day for MX to 600 mg/kg/day for BPA; the latter MED was approximately 600,000-fold higher than that for EE by the same route of exposure. The pooled dose responses for these chemicals are shown in Figure 1. This demonstrates the capability of the uterotrophic bioassay to detect estrogen agonists over a substantial concentration range.

The first article in this mini-monograph relates the results of the dose–response studies (Kanno et al. 2003b). The magnitude and shape of the dose–response curve for each of the five individual weak agonists was similar within a protocol. In several protocols,

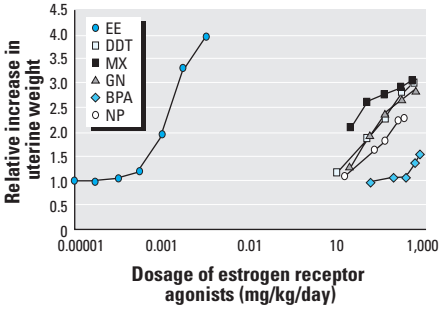

Figure 1. Mean uterine weight increase relative to vehicle control pooled across laboratories. Immature female rats were administered test substance doses by oral gavage for 3 consecutive days. The substances were EE, DDT, MX, GN , NP, and BPA.

no differences were observed in the MED, and in other protocols, the MED fell within a 3- to 4-fold concentration range despite differences in rat strains, diets, and other variables. No protocol or route of administration was clearly superior in the dose–response experiments. An additional observation from these MED data is that a blotted uterine weight increase of about 30–40% relative to controls results in statistical significance with a group size of six in a number of cases (Kanno et al. 2003b).

The second article relates the results of the coded dose studies (Kanno et al. 2003a). Here, the single doses selected were intended to be near the midpoint or in the upper half of a dose–response curve [ED_{50}–ED_{80} (effective dose concentration leading to approximately 50% of the maximum observed response in the uterine weight increase–effective dose concentrations leading to approximately 80% of the maximum observed response in the uterine weight increase)]. For some substances, however, the selected doses were actually near the MED in the low region of the dose–response curve. As this would place the dose in the low region of the dose–response curve, it would be anticipated that some laboratories would not achieve statistical significance. This, in fact, was the result (Table 1). The essential point is the dose dependency of a positive outcome. A positive chemical will consistently test positive only if a sufficient dose level is administered. As the dose is lowered, there will be a decreasing rate of positive results among laboratories until no statistically significant increase in uterine weight is achieved, a no-effect level across all laboratories. The other essential point is that there was reproducibility of the rate of positive detection at a given dose within the same protocol across laboratories (Table 1) when comparing the same administered dose to the MED calculated from the dose–response experiments.

Table 1. Immature and ovariectomized female rat comparing the MED and statistically significant results reported as positive identification/total number of trials.

Test substance	Administered dose by sc (mg/kg/day)	Immature female			Ovariectomized female		
		MED[a] (mg/kg/day)	Coded dose	Dose response	MED[a] (mg/kg/day)	Coded dose	Dose response
BPA	300	300	14/15	7/9	100	7/7	5/5
o,p´-DDT	100	200	6/16	1/4	100	4/7	1/3
GN	35	15	14/14	4/4	15	6/6	3/3
MX	500	100	15/15	4/4	100	6/6	3/3
NP	80	80	12/16	6/9	80	5/6	2/5

[a]MED given for pooled studies for a given protocol. Derived from dose–response experiments of Kanno et al. (2003b).

The third article relates the results of phytoestrogen analyses of the laboratory diets (Owens et al. 2003). The concern was that the responsiveness of the bioassay and its ability to detect weak agonists might be diminished by phytoestrogens, and all diets were found to contain phytoestrogens. An examination of the vehicle control weights and the responses to the weak agonists in different laboratories was made relative to an estimated dietary intake of phytoestrogens. The data indicated that no effect was evident on the responsiveness of the ovariectomized female to weak estrogen agonists. However, the data were suggestive of an effect for the immature rat model when GN intakes would exceed 50 mg/kg/day. This level is consistent with other toxicologic studies showing a lowest observable effect level in this range as well as the MED values in the dose–response studies. A review of rodent food consumption intakes revealed an approximate ratio for ovariectomized adult rats:immature rats:adult mice of 1:2:4 for food consumption on a kilogram of body weight basis. This would then lead to progressively higher phytoestrogens intakes. Although any conclusions must be drawn with caution until further more-controlled studies are done, experiments with immature rats or mice should consider limiting the dietary content of phytoestrogens to about 350 mg phytoestrogens/g diet and 175 mg phytoestrogens/g diet, respectively.

The fourth article relates the results of experiments addressing an animal husbandry variable, the housing of male and female animals in close proximity and the possible impact on female uterine weights (Ashby et al. 2003). In mice, housing adult males in close proximity to immature females can lead to premature puberty when the vomeronasal organs of the females are exposed to male urine. For the uterotrophic bioassay, this could mean that the response of these females would then be reduced. These studies directly tested this hypothesis in both rats and mice and found no effect on the uterotrophic bioassay.

Conclusions

This mini-monograph completes a successful, 4-year international validation program for the uterotrophic bioassay conducted under the auspices of the OECD. This success is due to the cooperation and sustained efforts of numerous scientists, participating laboratories, organizations, statisticians, and others. Thanks and congratulations are due to the many contributors.

The OECD validation studies demonstrate that the uterotrophic bioassay protocols are robust and reliable for identifying estrogen agonists and antagonists and are transferable across laboratories. Both the immature and ovariectomized rat protocols were highly reproducible among laboratories and were able to identify weak estrogen agonists, provided sufficient dose levels were administered. In conclusion, the rat uterotrophic bioassay, using either immature or ovariectomized animals, can be used as an effective screen for weak and potent estrogenic agonists.

References

Ashby J, Owens W, Odum J, Tinwell H. 2003. The intact immature rodent uterotrophic bioassay: possible effects on assay sensitivity of vomeronasal signals from male rodents and strain differences. Environ Health Perspect 111:1559–1567.

Chapin RE, Dulaney J, Wang Y, Lanning L, Davis B, Collins B, et al. 1999. The effects of 4-nonylphenol in rats: a multigeneration reproduction study. Toxicol Sci 52:80–91.

ECVAM (European Center for the Validation of Alternative Methods). 1995. Practical aspects of the validation of toxicity test procedures: the report and recommendations of ECVAM Workshop No. 5. ATLA 23:129–147.

ICCVAM (Intraagency Coordinating Committee on the Validation of Alternative Methods). 1997. Validation and Regulatory Acceptance of Toxicological Test Methods. A report of the Ad Hoc Interagency Coordinating Committee on the Validation of Alternative Methods. NIH Report No. 97-3981. Research Triangle Park, NC:National Institute of Environmental Health Sciences.

Kanno J, Onyon L, Haseman J, Fenner-Crisp P, Ashby J, Owens W. 2001. The OECD program to validate the rat uterotrophic bioassay to screen compounds for *in vivo* estrogenic responses: Phase 1. Environ Health Perspect 109:785–794.

Kanno J, Onyon L, Peddada S, Ashby J, Jacob E, Owens W. 2003a. The OECD program to validate the rat uterotrophic bioassay. Phase 2: coded single-dose studies. Environ Health Perspect 111:1550–1558.

———. 2003b. The OECD program to validate the rat uterotrophic bioassay. Phase 2: dose-response studies. Environ Health Perspect 111:1530–1549.

OECD (Organisation for Economic Co-operation and Development). 1998a. Report of the First Meeting of the OECD Endocrine Disrupter Testing and Assessment (EDTA) Working Group, 10–11 March 1998. ENV/MC/CHEM/RA(98)5. Paris:OECD.

———. 1998b. The Validation of Test Methods Considered for Adoption as OECD Test Guidelines. ENV/MC/CHEM(98)6. Paris:OECD.

———. 2001. Final Report of the Phase 1 of the Validation Study of the Uterotrophic Assay. [ENV/JM/TG/EDTA (2001)1/REV1]. Paris:OECD.

———. 2003. Detailed Background Review of the Uterotrophic Bioassay: Summary of the Available Literature in Support of the Project of the OECD Task Force on Endocrine Disrupters Testing and Assessment (EDTA) to Standardise and Validate the Uterotrophic Bioassay. OECD Environment, Health and Safety Publication Series on Testing and Assessment, No. 38. ENV/JM/MONO(2003)1. Paris:OECD.

Owens JW, Ashby J. 2002. Critical review and evaluation of the uterotrophic bioassay for the identification of possible estrogen agonists and antagonists: in support of the validation of the OECD uterotrophic protocols for the laboratory rodent. Crit Rev Toxicol 32:445–520.

Owens W, Ashby J, Odum J, Onyon L. 2003. The OECD program to validate the rat uterotrophic bioassay. Phase 2: dietary phytoestrogen analyses. Environ Health Perspect 111:1559–1567.

Tyl RW, Myers CB, Marr MC, Thomas BF, Keimowitz AR, Brine DR, et al. 2002. Three-generation reproductive toxicity evaluation of bisphenol A in CD Sprague-Dawley rats. Toxicol Sci 68:121–146.